User-Centric Health Data Using Self-sovereign Identities*

Alexandre Siqueira[1], **Arlindo Flavio da Conceição**[1], **Vladimir Rocha**[2]

[1] Universidade Federal de São Paulo (UNIFESP)

[2]Universidade Federal do ABC (UFABC)

{alexandre.siqueira, arlindo.conceicao}@unifesp.br,

vladimir.rocha@ufabc.edu.br

Abstract. *This article presents the potential use of the Self-Sovereign Identities (SSI), combining with Distributed Ledger Technologies (DLT), to improve the privacy and control of health data. The paper presents the SSI technology, lists the prominent use cases of decentralized identities in the health area, and discusses an effective blockchain-based architecture. The main contributions of the article are: (i) mapping SSI general and abstract concepts, e.g., issuers and holders, to the health domain concepts, e.g., physicians and patients; (ii) presenting and instantiating an architecture to deal with the use cases mentioned, effectively organizing the data in a user-centric way, that uses well-known SSI and Blockchain technologies.*

1. Introduction

Personal health records reveal a conflict between privacy protection laws and the ability to share personal information. Throughout life, people interact with doctors and healthcare service providers countless times, either for routine appointments or medical treatment. In these interactions, clinical information is obtained that could constitute a significant batch of medical records for future use. However, a single patient might have his/her medical data spread across several healthcare service providers, and thus create siloed databases that, in the current state — fragmented, are useless for clinical application outside those silos. Added to this scenario, the increasing adoption of wearable devices, for collecting information about health and lifestyle, is creating a new patient-generated health data silo. Thus, new methods are needed for creating a decentralized data structure, controlled by the users, to maintain their medical records. This appears to be the pathway to enabling health providers to store medical information while granting data ownership to the rightful owners — the patients.

In health domain, the problem of data ownership can be overcome by the use of Self-sovereign Identities (SSI). SSI are designed to enable entities to control its DIDs, using cryptographic tools, such as digital signatures [López 2020]. With that, users in the health domain could use these identifiers to interact among them (exchanging their identities), in a secure and reliable way, without loosing the data ownership. The use of this kind of technology and identities could be a means of confronting the challenge of sharing and securing sensitive medical information among healthcare parties, as well as ensuring patients maintain sovereignty over their data.

*This research is part of the INCT of the Future Internet for Smart Cities funded by CNPq, proc. 465446/2014-0, CAPES proc. 88887.136422/2017-00, and FAPESP, proc. 2014/50937-1.

In the remaining of this article, Section 2 discusses related work. Section 3 presents the principles of SSI. Section 4 proposes the architecture based on SSI and DLT frameworks. Section 5 discusses the practical open problems related to the application of SSI in healthcare. Finally, in Section 6 we present ideas for future works and our conclusions.

2. Related work

In recent years, several authors explored the idea of using intelligent agents in healthcare context to provide interoperability [Isern and Moreno 2016, Barrué et al. 2015, Wimmer 2014]. More recently, some authors explored blockchain technology [Shrier et al. 2016, Liu 2016, da Conceição et al. 2018], most of them proposing strategies to improve mobility and security in EHRs using distributed ledgers.

Regarding using both SSI and DLT in the healthcare context, to the best of our knowledge, we found few works directly related to ours, two of them discussed below.

Bouras *et al*. [Bouras et al. 2020] discuss several aspects to achieve a decentralized identity management, using blockchain, in health systems. Among those aspects, they identify the key players, their roles, and the different health scenarios in which SSI and DLT technologies can be applied. In these scenarios, the authors identify the players associated and the requirements necessary (such as interoperability, data standards, anonymity, scalability, etc.) to deploy them in the real world. Besides, the work also lists several decentralized identity management solutions that use the SSI and DLT technologies, but only one of them (Evernym) has the Healthcare domain as an industry target.

Houtan *et al*. [Houtan et al. 2020] as well as the aforementioned work, also discuss on how the SSI and Blockchain can be applied in the health domain. The difference is in the focus on the challenges that arise when creating a healthcare information system (HIS) that must be decentralized, private, and secure in order to exchange patients' data. Among the challenges in using DLTs, the authors identify that the blockchain category (e.g., public, private, hybrid), smart contract programming language, and consensus protocol create trade-off that must be analyzed by the architects and designers in order to avoid functional and non-functional misbehavior or security problems. This work presented and analyzed several solutions using blockchain as the core for creating HIS, but concludes that most of them are not ready to be deployed in real-world scenarios.

3. Self-Sovereign Identity (SSI)

A person walks into a bar and asks for a drink. The bartender then asks to show some identification to confirm that he/she is old enough to buy liquor. The person presents his/her state-issued driver's license, containing the name, photo, date of birth, and other personal information. The bartender recognizes the presented document, notices the resemblance between the photo and the person, figures the persons' age by his date of birth, and, finally, confirms that the person can place the order.

Ordinary interactions like this, proposed by Windley [ComputerWorld 2018], are the inspiration behind a concept called self-sovereign identity (SSI). The SSI model can be categorized as a user-centric model because it grants the users control over their data. User-centric approaches for managing digital identities date back to 2005 when

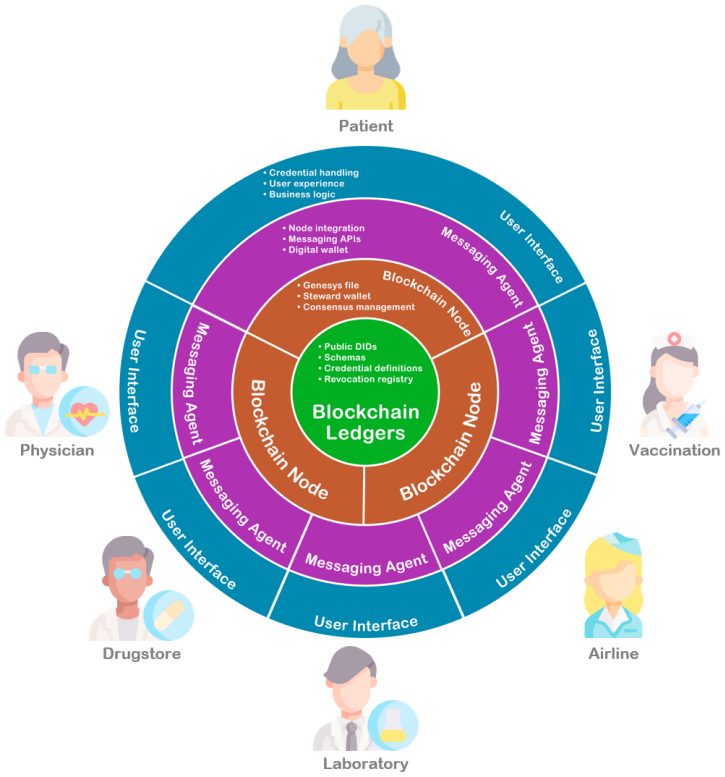

Figure 1. Context diagram representing layers and responsibilities

Josang described a model where users would store their identifiers and credentials in a portable hardware device secured by a local secret, like a PIN [Jøsang and Pope 2005]. The advancements in portable devices and distributed ledger technologies brought improvements to the user-centric model, enabling the creation of two essential elements for SSI [López 2020]:

- **Decentralized ledgers:** distributed records structures that store cryptographic proofs such as digital signatures and timestamps, allowing anyone to verify digital credentials issued by entities without the need of a central source of trust.
- **Digital wallets:** portable and secure personal repositories that allow users to manage identities and verifiable credentials within the phones, completely protected and under their control [López 2020]. It also enables voluntary information disclosure situations: users can select what information to disclose to whom.

4. Architecture

This section presents the proposed architecture of the blockchain-based self-sovereign health registry system. Figure 1 portrays a context view, where the concentric circles depict the layers of the system.

The two inner layers are the foundation of the architecture. They represent the blockchain network and are common to all actors in the system. The innermost layer describes the **blockchain ledgers**, public information repositories that store data using a DLT.

The layer that surrounds the ledgers represents the **blockchain nodes**. Those

nodes are the machines that support the blockchain network, managing the data that is stored in the ledgers. The **consensus management** algorithms ensure that every node has a synchronized copy of the ledgers. There can be as many nodes as needed to keep the blockchain network running efficiently, but the initial nodes are known as "Genesis nodes". Those initial nodes provide a unique identification of the network and are described in a file called **Genesis file**. The nodes and the messaging agents (explained below) must hold this file to connect to the network. The nodes are responsible for writing and reading data from the ledgers, and therefore, they need their own DIDs. Part of the node's DIDs is stored in their digital wallets, named **steward wallets**.

The **messaging agent** layer is responsible for holding **digital wallets** encrypted and protected, and for connecting with others blockchain nodes and agents. These two different connections are essential for understanding how information flows among entities and are explained in detail in the following sections. In summary, those connections can be described:

- **Node communication:** using HTTP protocol, the agents connect to the blockchain network to write and read public data from the ledgers. The agents also expose **messaging APIs** to user interface components for credential handling.
- **DIDComm:** a protocol for secure peer-to-peer communication between agents, DIDComm allows entity agents to establish trusted relationships and exchange private messages and credentials.

The outermost layer represents the **user interface**, where specific **business logic** is implemented. The user interface provides entity users with **credential handling** capabilities, such as credential issuance, credential verification, and relationship establishment. In straight terms, the user interface layer allows users to interact with the system.

4.1. Components and adopted technologies

The proposed architecture is comprised of several software components. Figure 2 depicts a diagram that groups these components in software packages and distributes them in the architectural layers shown in Figure 1, represented as stacked blocks.

The **Blockchain ledgers layer** uses Hyperleger Indy's Plenum module to handle the storage of public identity information in distributed ledgers and Hyperledger Ursa's library to perform the necessary cryptographic primitives. Hyperledger Indy and Hyperledger Aries also reference Ursa's crypto library to reuse complex cryptographic primitives across Hyperledger projects. The **Blockchain node layer** runs over Hyperledger Indy's Node, Plenum and SDK modules, managing the blockchain network and its consensus mechanisms.

The **Messaging agent layer** represents the clients of the Blockchain network and runs over the Hyperledger Aries agent and embedded Indy SDK module. The **User interface layer** is an abstraction for the healthcare software that uses Messaging agent APIs to provide specific business journeys for Issuers, Holders, and Verifiers.

5. Discussion

We implemented the proposed architecture for validation, using the Indy/Aries architecture, installed on a cloud server, and the Trinsic wallet[1].

[1] `https://trinsic.id`

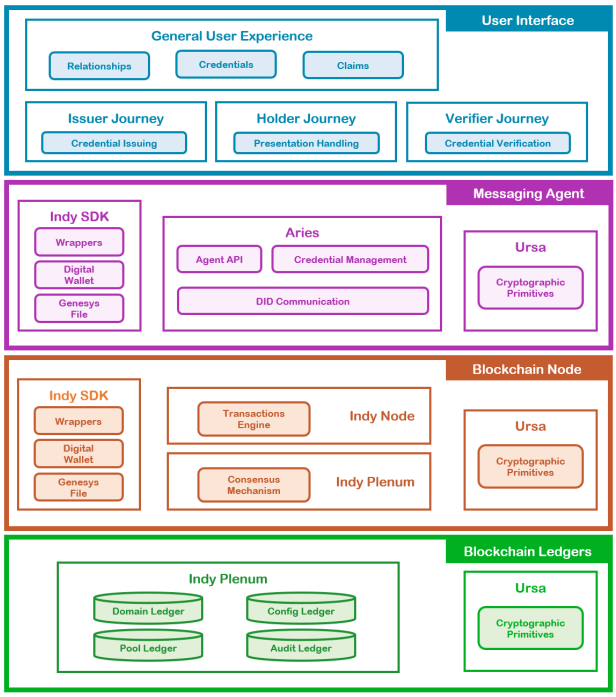

Figure 2. Architectural layers and stacked software packages

The main paradigm shift in system design is the ownership of data, which are under the control of the patient. This reduces the amount of data held by third parties, therefore reducing the potential for data leakage. However, it is possible to highlight new challenges.

The first challenge is **client-side storage**. In an ideal scenario, this storage should be in a local repository (e.g., patient's cell phone) with automatic backup to the cloud, similar to how Dropbox currently operates. The techniques to manage secure local writing and automatic cloud backup must be developed. In addition, we have to find out whether the storage medium on a modern cell phone is adequate for storing basic health data. Of course, some data, such as images and videos, should only be stored in the cloud.

Another challenge is the **usability** of the system. We believe that soon, the population will use services based on reading QRCodes and digital signatures. Still, the ease of using these interfaces should also be a determining factor for its wide adoption. An important line of research is to make the exchange of health data between Issuer and Holder automatic — or almost automatic, especially when data collection takes place through sensors of Internet of Things.

A third challenge is the **maintenance** of the computing infrastructure. It is clear that patients must maintain local nodes, and Issuers and Verifiers must maintain messaging agents. But who should support the blockchain Indy nodes? The business model is not well defined.

Finally, we believe there is a significant economic potential in developing software, as each entity, Issuer or Verifier, may need to customize their messaging agents to meet their business. We also believe in the emergence of new business chains based on the trustworthiness of credentials and data availability. For example, a heart test could be

used to enroll in a gym; attendance at the gym can generate discounts on health insurance; health coverage can create discounts on life insurance, and so on.

6. Conclusions

This work exploits the main applications of SSI in healthcare scenarios. In our proposal, all data is owned by the patients. And it relies on SSI, blockchain technology and well-known standards.

The architecture is flexible in order to satisfy several standardized uses cases. It does not solve dilemmas related to data accessibility versus privacy, but it provides appropriate means for the explicit consideration of each of these issues.

For the future, we plan to implement a functional prototype of the proposed architecture, shown in Section 4, and obtain feedback from patients and healthcare professionals and institutions about the usability of the system. In addition, after to validate the basic solution, we could create more complex data health management models.

References

Barrué, C. et al. (2015). Using multi-agent systems to mediate in an assistive social network for elder population. In *Proceedings of the 18th International Conference of the Catalan Association for Artificial Intelligence*, volume 277, page 120.

Bouras, M. A., Lu, Q., Zhang, F., Wan, Y., Zhang, T., and Ning, H. (2020). Distributed Ledger Technology for eHealth Identity Privacy: State of The Art and Future Perspective. *Sensors*, 20(2).

ComputerWorld (2018). How blockchain makes self-sovereign identities possible. Disponível em: `https://www.computerworld.com/article/3244128/how-blockchain-makes-self-sovereign-identities-possible.html`. Último Acesso em: 06.04.2021.

da Conceição, A. F., da Silva, F. S. C., Rocha, V., Locoro, A., and Barguil, J. M. (2018). Electronic health records using blockchain technology. In *WBlockchain, in conjunction with Simpósio Brasileiro de Redes de Computadores (SBRC)*.

Houtan, B., Hafid, A. S., and Makrakis, D. (2020). A survey on blockchain-based self-sovereign patient identity in healthcare. *IEEE Access*, 8:90478–90494.

Isern, D. and Moreno, A. (2016). A systematic literature review of agents applied in healthcare. *Journal of medical systems*, 40(2):43.

Jøsang, A. and Pope, S. (2005). User centric identity management. In *AusCERT Asia Pacific information technology security conference*, page 77.

Liu, P. T. S. (2016). Medical record system using blockchain, big data and tokenization. In *Information and Communications Security*, pages 254–261. Springer.

López, M. A. (2020). Self-sovereign identity: The future of identity: Self-sovereignity, digital wallets, and blockchain. Último Acesso em: 11.04.2021.

Shrier, A. A. et al. (2016). Office of the national coordinator for health information technology us department of health and human services.

Wimmer, H. (2014). A multi-agent system for healthcare data privacy. In *AMCIS*.